Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 133 (2018) 1066–1073

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Robotics and Smart Manufacturing (RoSMa2018)

Implementation of a Stereo vision based system for visual feedback control of Robotic Arm for space manipulations

Sangeetha G.R[a], Nishank Kumar[a], Hari P.R[a] , Sasikumar S[a]

[a]ISRO Inertial systems Unit, Trivandrum

Abstract

The paper discusses the design and development of a stereo vision based system for use in robotic arms suitable for space applications. Based on the co-ordinates of a target estimated using images of the scene, a robotic arm can be used to precisely position to the target. This will find application during sample collections in interplanetary missions.
The paper discusses the visual servo loop including the stereo vision algorithms which are used to determine the co-ordinates of the target. , derivation of DH parameters and the development of inverse kinematics matrix and solutions for a proto three degree of freedom robotic arm on which the stereo vision system was tested. The prototype robotic arm reaches to a target to well within 2cm.

Keywords:"space robotics, stereo vision, disparity map, kinematics"

1. Introduction

As India is embarking on its interplanetary odyssey, primarily aimed at conducting scientific experiments and observations, a large number of suitable scientific payloads need to be designed and developed. Being a hitherto unknown territory, visual mapping, sample collection and analysis will greatly play a major role in defining and enhancing the value of these expeditions. Space robotics augmented with visual servo control is an area which can find great applications to meet the above requirement.

In the first time-lesser known interplanetary environments, navigation can be done by different methods like LiDARs (Light Detection and Ranging), vision based servoing etc. Vision is a powerful tool for navigation as is the case with man himself. Visual information is used to locate and identify things, and to decide how they should be grabbed. Visual feedback is then used to guide the hand to the target. In a similar manner, machine vision can be used to co-ordinate a robotic arm. The three dimensional information about the environment is essential for robot movement and object inspection in robotic applications. One method of obtaining depth information is by using a stereo vision

10.1016/j.procs.2018.07.031

system. Though LiDAR requires much less processing, it is relatively expensive and requires active laser projection, when compared to stereo vision which is cheaper, completely passive, but computationally moderate in comparison. Stereo vision is the method used to determine the three dimensional information of a scene from two simultaneous images of the scene [1].

This paper discusses the design and development of a stereo vision system which is based on a look-and-move hierarchical approach, where the acquired image is used to direct the manipulator to any specific location and manipulate objects. The scheme has been tested on a three-degrees-of-freedom robot manipulator with a stereo camera mounted on its end-effector and it can position itself automatically to the required position with respect to a target object [2]. The kinematic modelling of the arm is also discussed here as it forms the part of the end to end system. The robot manipulator system attached to a mobile rover can be used in future unmanned moon or mars mission to explore the environment.

2. Stereo Vision based system for Robotic arm manipulations

The basic block diagram of the stereo vision system in use in the robotic arm is as shown in figure 1.The robotic arm can be used to precisely acquire a target by estimating its co-ordinates using real time acquired images of the scene. Using the stereo camera mounted on the base of the Robotic arm, the left and right view images of the scene are taken. The rectified images are then processed using stereo vision algorithms to obtain the disparity map of the scene .With human intervention, a suitable target object is chosen and the real world coordinates of the target are obtained from the computed disparity. The arm takes the (x,y,z) coordinates of the object as its input and generates the desired coordinates of its joint space via the inverse kinematics solution. The joint encoders helps to know the current position of the arm and actuators move it to the desired position, thus achieving the target co-ordinates in real world space. The robotic arm thus consists of a two tier control system- an outer loop that takes the visual cues as feedback and an inner low level driver/controller at the joints that actuates the arm based on feedback from encoder.

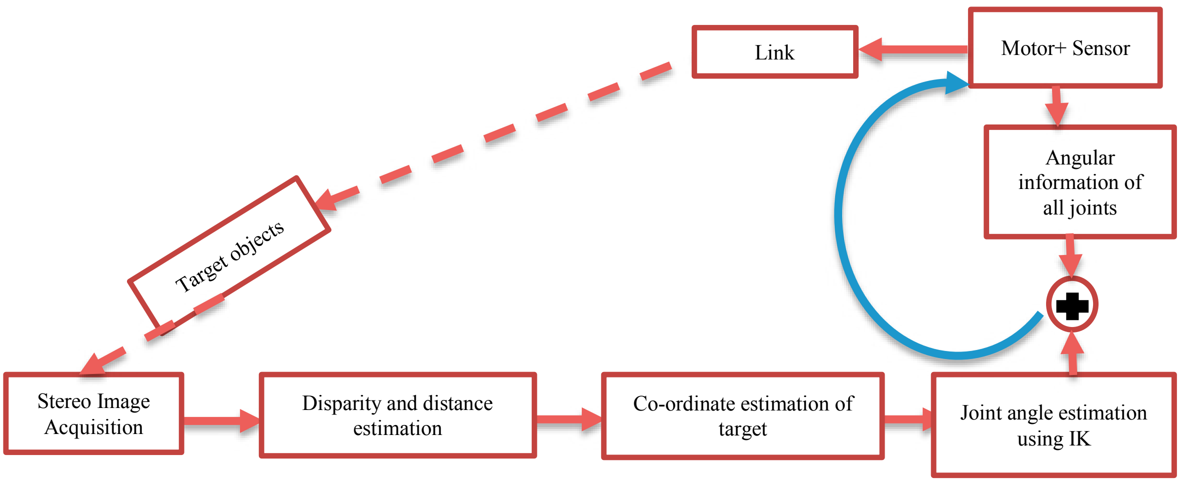


Fig1. Block diagram of Robotic Arm with stereo vision system

3. Distance measurement using stereo triangulation method

Stereo vision system has two images-right and left at its disposal. If by a means, two matching pixels can be found in the two images, then the horizontal distance between the matching pixels is defined as disparity between the two. Thus for a set of two images, the disparity can be computed for any number of matching pix els which gives a disparity map .Once the disparity is known, the distance can be determined as in equation 1.

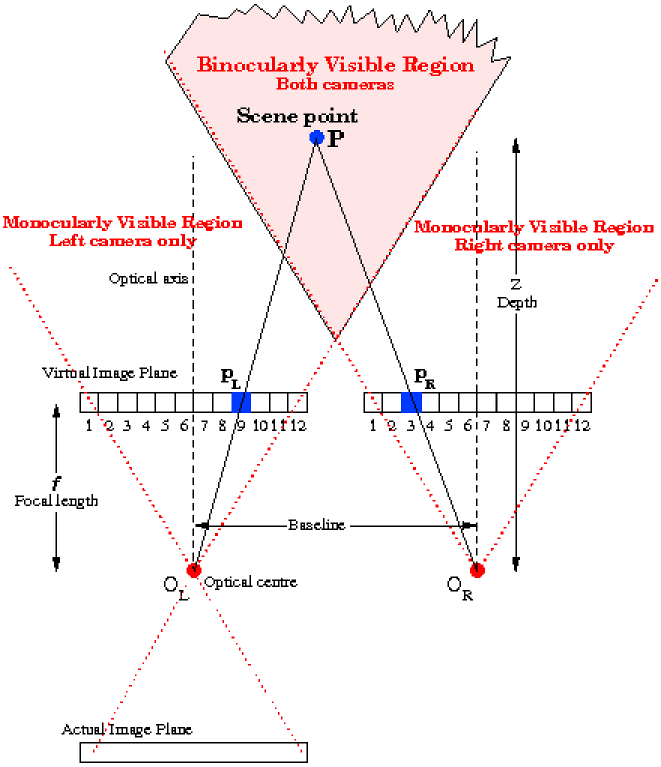


Fig 2. Stereo triangulation method [3]

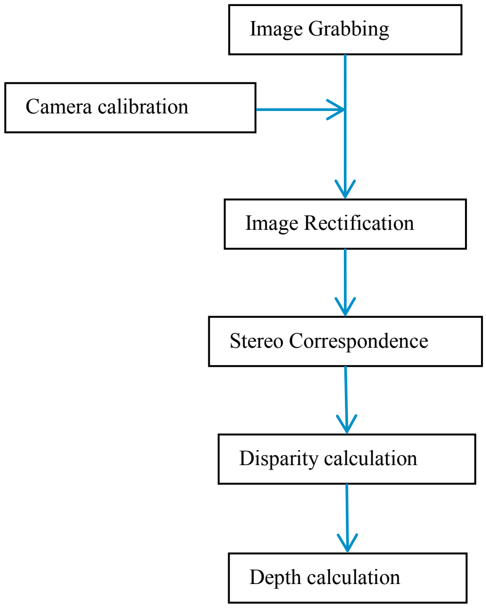


Fig3. Steps in estimation of depth from stereo images

$$\boldsymbol{Z = \frac{f.b}{x_l - x_r} = \frac{f.b}{pd}} \qquad (1)$$

Where d is the disparity, p is the unit pixel width, f is the focal length of the camera, and b is the baseline as shown in figure 2.

3.1. *Steps in development of a stereo vision system:* The basic steps in the estimation of distance using stereo vision system is explained next and is shown in figure 3.

3.1.1 *Image grabbing*: The system for calculating depth starts with grabbing the images. The stereo vision camera system consists of two cameras mounted on a single fixture displaced horizontally from each other by a fixed distance called the baseline. Stereovision cameras thus capture left and right view images of the same scene. In this work, Bumblebee stereo vision camera system BB2-03S2 from Point Grey is used [4].

3.1.2 *Camera calibration*: It is the process of obtaining the intrinsic and extrinsic parameters of the camera by using a calibration object. The calibration and estimation of camera parameters are done for Bumblebee Point Grey stereo vision camera BB2-03S2. The pixel width is estimated as *9.34um and Focal length as 0.295m*

3.1.3 *Stereo image rectification*: This technique projects images onto a common image plane in such a way that the corresponding points have the same row coordinates, thus enabling search along the horizontal lines of the rectified images. Thus it enables 1-D search and reduces the computational time and complexity in real time applications. In this current set of stereo vision system, rectification can be done away with since the

translational and the rotation between the two cameras forming the single stereo vision system is arrested and is known.

3.1.4 Stereo correspondence problem:

To determine the disparity, a process of matching areas and or specific locations between the two stereo cameras is essential and this is the solution of the stereo correspondence problem.

There are both global and local algorithms for determining the correspondence [5]. Area-based local methods are less computationally expensive than globally optimized pixel-based correspondence. Hence they are preferred for real time execution.

Sum of absolute differences algorithm (SAD) is the cost function used in this work. It is selected after a comparative study on various stereo correspondence algorithms and SAD emerged the winner in terms of computational time and memory which are essential parameters for real time algorithms to be implemented onboard. In SAD, the correspondence problem is approached by comparing blocks of pixels in each image. The dissimilarity is the sum of differences in pixel intensities. The match with the highest similarity wins.

Let f(i,j)be the intensity of the pixel at coordinates (i, j) in the reference image. For the second image g(i, j) is used. N defines the extend of the block in either direction around the center pixel, the block has therefore width and height (2N+1).The variable d is the disparity, defined here as the offset of the block along the epipolar line in the second image.

$$\mathrm{SAD}(x, y, d) = \sum_{i=x-N}^{x+N} \sum_{j=y-N}^{y+N} |f_{i,j} - g_{i+d,j}|$$

SAD is computed for all possible disparities in the disparity range, D, and the disparity, associated with the minimum SAD value is selected for each pixel (x, y).

$$\mathrm{disparity}(x, y) = \min_{d \in D} \left(\mathrm{sad}(x + d, y)\right) \quad (2)$$

When applied to a rectified image, the disparity space image containing the disparity for each pixel along with the camera parameters was used to find the distance using the stereo triangulation method. The algorithm has been implemented in MATLAB. Pre -processing is done to account for photometric variations between cameras in stereo rig. Image normalization is done to have local statistics with zero mean and unity standard deviation. The LoG filter is used to suppress low frequency background intensity and very high frequency noise. Once the disparity map is obtained as in figure 4, the (x,y,z) coordinates of any point in the scene can be estimated as shown in fig 5.

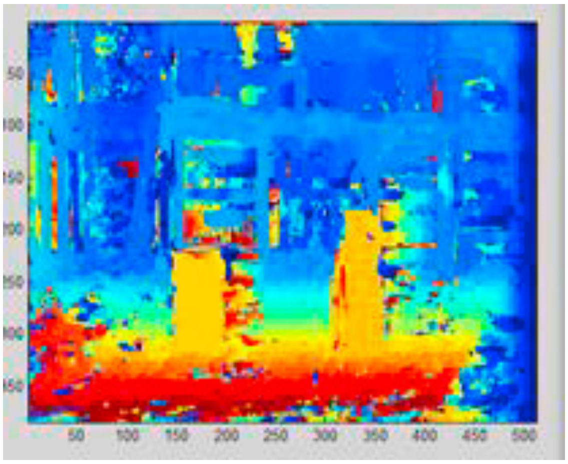


Fig4. Disparity map obtained for a real world scene

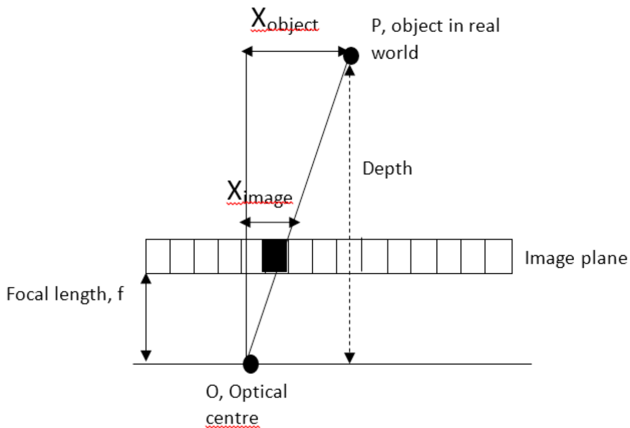


Fig5. Estimation of real world co- ordinates from image co ordinates

3.1.5 Estimation of (x,y,z) co ordinates :

By similar triangles,

$$X_{object} = \frac{X_{image} * \text{Disparity} * \text{pixel width}}{\text{focal length}}$$

Depth = Disparity in pixels *pixel width.

4. Development of Arm Kinematics

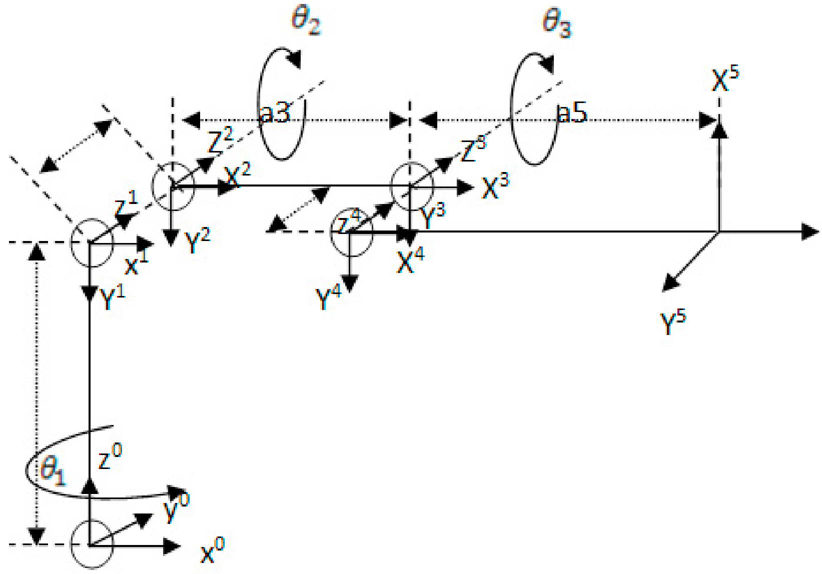


Fig 6. The DH representation of the developed robotic arm

Once the target co- ordinates are estimated from the stereo vision system, the arm should be directed to reach the coordinates in space. For this purpose a kinematic modelling of the arm should be done. To test the stereo vision loop, a three DOF articulated structure with revolute joints is used.

4.1. Direct kinematics:

The direct kinematics refers to the mapping of joint position vector Q to the position vector X of the robot end-effector i.e.$X = f(Q)$.It finds the current position and orientation of the tool tip given the joint angles. First the link coordinate diagram is made and based on it through a series of coordinate transformations we can arrive at the homogeneous transformation matrix of dimension 4X1, which gives the position and orientation information, in terms of joint angles.

Link coordinates: The arm can be modeled as a chain of rigid links interconnected by revolute joints. The various related parameters are the joint (joint angle-θ, joint distance -d) and the link parameters (link twist angle- α, link length -a) as shown below in table1. The values of the joint and link parameters for the developed arm are d1=0.58m;d2=0.045m;a3=0.305m;d4=-0.055m;a5=0.676m;

Table 1. DH parameters for the current arm

Sl. No	θ	d	a	α
1	θ_1	d_1	0	-90°
2	0	d_2	0	0
3	θ_3	0	a3	0
4	0	d_4	0	0
5	θ_5	0	a5	-90°

Link coordinates are fixed as per the DH representation guidelines [6] and are as shown in figure 6.
The arm equation is derived by a series of coordinate transformations including translation and rotations from base to tool tip. The position and orientation of the tool frame relative to the base frame is found by considering the 3 consecutive link transformations matrices relating fixed to adjacent links.

$$\mathrm{T}_{\mathrm{base}}^{\mathrm{wrist}} = \mathrm{T}_{\mathrm{base}}^{\mathrm{shoulder}} * \mathrm{T}_{\mathrm{shoulder}}^{\mathrm{elbow}} * \mathrm{T}_{\mathrm{elbow}}^{\mathrm{wrist}}$$

By suitable combinations, we can get the 4X4 homogeneous transformation matrix from which the tool configuration vector of 3X1 is reproduced below which gives the (x,y,z)co-ordinates of the tool tip with respect to the base co-ordinates, in equation 3

$$\begin{bmatrix} x_0^5 \\ y_0^5 \\ z_0^5 \end{bmatrix} = \begin{vmatrix} a_5\cos\theta_1\cos\theta_2\cos\theta_3 - a_5\cos\theta_1\sin\theta_2\sin\theta_3 - d_4\sin\theta_1 - a_3\cos\theta_1\cos\theta_2 - d_2\sin\theta_1 \\ a_5\sin\theta_1\cos\theta_2\cos\theta_3 - a_5\sin\theta_1\sin\theta_2\sin\theta_3 + d_4\cos\theta_1 - a_3\sin\theta_1\cos\theta_2 + d_2\cos\theta_1 \\ a_5\sin\theta_2\cos\theta_3 + a_5\cos\theta_2\sin\theta_3 - a_3\sin\theta_2 + d_1 \end{vmatrix} \tag{3}$$

(ii) Inverse Kinematics
Inverse kinematics solution gives the joint angles, $\theta_1, \theta_2, \theta_3$ necessary to reach the desired [x,y,z] co ordinates of the target. The three joint angles can be obtained analytically as follows

$$T = \begin{bmatrix} R & D \\ 0 & 1 \end{bmatrix}$$

$$R - orientation\ of\ end\ effector\ w.r.t\ base$$

$$D - position\ of\ end\ w.r.t\ base$$

The entry at D forms the tool position vector giving the relative position of end effector w.r. t base. From the tool configuration vector, using trigonometric identities, the joint angles are obtained as given below.

$$\theta_3 = \mathrm{atan2}(\sin\theta_j, \cos\theta_j) \tag{ㄷㅈ}$$

$$\theta_2 = \mathrm{atan2}(rd - sc, rc + sd) \tag{ㄷㅊ}$$

Where $r = a_5\cos\theta_3 + a_3$, $s = a_5\sin\theta_3$

$$c = a_5\cos(\theta_2 + \theta_3) + a_3\cos\theta_2$$

$$d = a_5\sin(\theta_2 + \theta_3) + a_3\sin\theta_2$$

And $$\theta_1 = \mathrm{atan2}(My - Nx, Mx + Ny) \tag{ㄷㅌ}$$

Where $N = d_4 + d_2, M = a_5\cos(\theta_2 + \theta_3) + a_3\cos\theta_3$

5. Testing And Verification of Stereo Vision System

The stereo vision system was tested by integrating it with the prototype 3 DOF robotic arm. The kinematics algorithms were implemented in MATLAB^{R} 2012b and the interface used was NI6259 DAQ PCIcard. A graphical user interface is developed in the MATLAB environment as shown in fig 7. At Power on, the arm can be commanded to any desired home position. The GUI reads the image obtained from the bumblebee camera. The depth map of the scene is generated by the SAD algorithm.

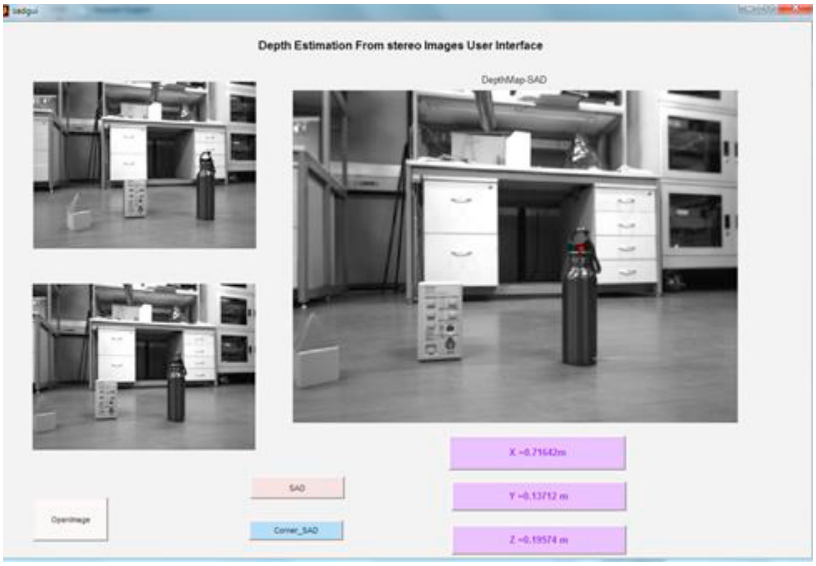


Fig 7. Graphical user interface for stereo vision system

The desired object in the scene may be chosen by human intervention and the real world coordinates of the target are obtained by computation from the depth map. With a knowledge of the desired co-ordinates obtained from the stereo algorithm, the inverse kinematics algorithm gives the angles to be attained by the three joints. The joints are then actuated based on the feedback from the encoders using stepper motor through amplifiers and drive circuit.

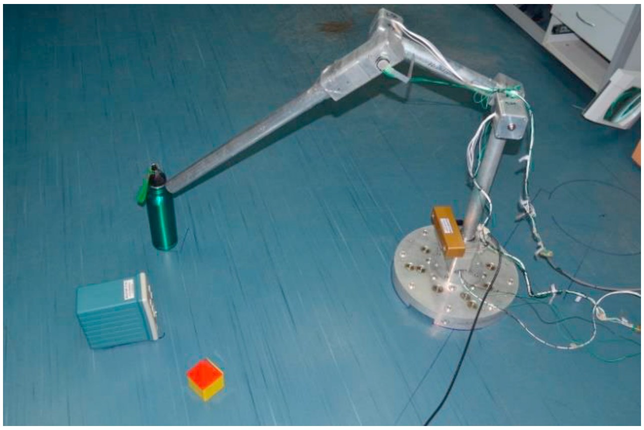


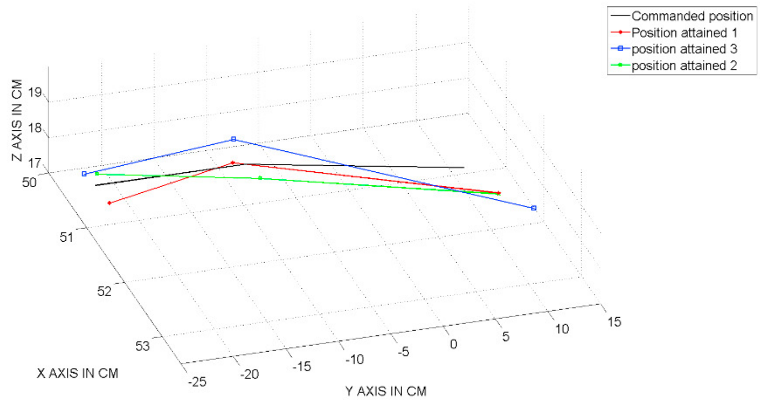
Fig 8. The stereo vision based robotic system reaching a target object in real world

To verify the performance of the ARM in reaching the chosen targets in a scene, objects were placed in known locations. The stereo vision based robotic arm system was used to estimate the x, y and z co-ordinates of the target objects and move to the chosen targets as shown in figure 8. The exercise was done from 50cm to 75 cm over three different ranges and over a span of 3 different objects for representative points over the entire range of the arm.
It is seen that the arm reaches the target with a best achieved accuracy of 2 cm as shown in fig 9.
The accuracy can be improved further by using a higher resolution imaging sensor as well by employing algorithms to estimate disparity with sub pixel accuracy.

6. Conclusion And Future Scope

Stereo vision based system has been developed and implemented, which is capable of obtaining the depth information

of objects in a real world scene from the 2D image of the scene. The vision system is integrated with the robotic arm and it can position to a target object in real world within a best achieved accuracy of 2cm. This is comparable to the values achieved by major positioning systems used for space manipulations. The stereo vision algorithms have been chosen to enable real time execution with lesser computational complexity in terms of memory and time consumption.



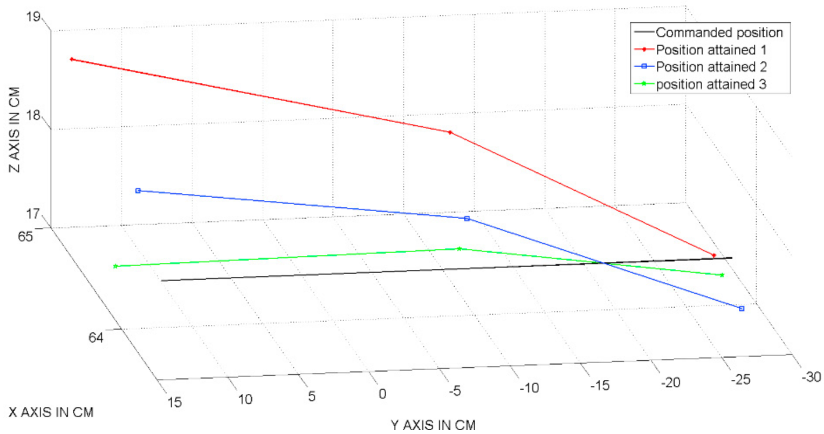


Fig 9. Plot showing the position accuracy achieved by the stereo vision based robotic arm over the work space.

REFERENCE

[1] Ueno, Haruki, and Yukio Saito. "Model-based vision and intelligent task scheduling for autonomous human-type robot arm." Robotics and autonomous systems 18.1(1996): 195-206.

[2] Proto model of 3DOF Robotic Arm with stereovision based pointing, IISU/TR/2009/2016 (Internal Technical report, ISRO Inertial systems Unit)

[3] Stereo vision, lectures by Mircea Nicolescu, University of Nevada at Reno.

[4] BB2-03S2, Bumblebee specifications, Bumblebee stereo vision camera systems

[5] Scharstein, Daniel, and Richard Szeliski. "A taxonomy and evaluation of dense twoframe stereo correspondence algorithms." International journal of computer vision 47.1-3 (2002): 7-42.

[6] Mark W. Spong, Robot modelling and control